LIBRARY
OF THE
MASSACHUSETTS INSTITUTE
OF TECHNOLOGY

WORKING PAPER
ALFRED P. SLOAN SCHOOL OF MANAGEMENT

Relaxation Methods for Pure and Mixed Integer Programming Problems*

by

G.A. Gorry, J.F. Shapiro, L.A. Wolsey

April, 1970

WP 456-70

MASSACHUSETTS
INSTITUTE OF TECHNOLOGY
50 MEMORIAL DRIVE
CAMBRIDGE, MASSACHUSETTS 02139

MASS. INST. TECH.
MAY 6 1970
DEWEY LIBRARY

Relaxation Methods for Pure and Mixed Integer Programming Problems*

by

G.A. Gorry, J.F. Shapiro, L.A. Wolsey

April, 1970

WP 456-70

*This research was supported in part by the U.S. Army Research Office (Durham).

Abstract

The usefulness of group theoretic methods in solving integer programming (IP) problems is extended by procedures for controlling the size of the groups. The main procedure given shows how an optimal linear programming basis can be altered to reduce the magnitude of its determinant thereby reducing the size of the group induced by the basis. An adoption of Bender's mixed IP algorithm is given which uses these methods. Some limited computational experience is given.

535674

1. Introduction

Computational experience with a group theoretic integer programming (IP) algorithm has been unusually promising; see reference 6. A major drawback, however, of this method in the past has been the size of the finite abelian groups encountered during the computation of certain IP problems. In this paper, we show how the difficulty can be overcome by the application of some rather simple number theoretic procedures. Although these procedures cannot be guaranteed to reduce the size of the induced groups, computational experience to date has shown them to be efficacious in practically all cases. The central idea is that a given IP problem can be usefully transformed by dividing all the coefficients of an inequality by a rational divisor greater than 1, and then taking integer parts. Such a transformation of the problem, called a relaxation, has the property that all feasible solutions to the original IP problem are feasible in the transformed or relaxed problem. The idea for this particular relaxation method in IP is originally due to Wolsey [12]. The term relaxation is also used by Geoffrion in [3] to describe mathematical programming methods conceived in the same spirit as the ones here.

The plan of the paper is the following. Section 2 consists mainly of definitions and ideas from [12] for changing the coefficients of a given IP problem by relaxation and other means. The following section demonstrates that relaxation of constraints can be made as required after the given IP problem is solved as a linear programming (LP) problem. Section 4 is concerned with a generalized Lagrange multiplier technique for "pricing out" the inaccuracies introduced by the procedures of sections 2 and 3. Section 5 contains an adaption of Bender's method [1] for mixed IP in which group theoretic methods are used to solve a series of pure IP

subproblems. The relaxation procedures of the previous sections are necessary for the use of group theoretic methods on the pure IP subproblems. The final section gives some computational experience.

2. Relaxation and Restriction of IP Problems

Consider the two optimization problems

$$\begin{aligned} &\min\ f(x) \\ &\text{s.t. } x \in X_1, \end{aligned} \tag{2.1}$$

and

$$\begin{aligned} &\min\ f(x) \\ &\text{s.t. } x \in X_2, \end{aligned} \tag{2.2}$$

where X_1 and X_2 are subsets of E^n. Problem (2.2) is said to be a relaxation (restriction) of problem (2.1) if $X_1 \subset X_2$ ($X_2 \subset X_1$). Problems (2.1) and (2.2) are equivalent if $X_1 = X_2$.

Consider now the IP problem

$$\begin{aligned} &\min \sum_{j=1}^{n} c_j x_j \\ &\text{s.t. } \sum_{j=1}^{n} a_{ij} x_j \leq b_i \qquad i = 1,\ldots,m \\ &\qquad x_j \text{ non-negative integer, } j = 1,\ldots,n, \end{aligned} \tag{2.3}$$

where all the coefficients are integer. It is convenient to define the 1 x n vector $c = (c_j)$, the m x n matrix $A = (a_{ij})$, and the m x 1 vector $b = (b_i)$.

Let $\lambda_i \geq 1$ be a rational divisor selected for row i, i = 1,...,m. It is shown in [12] that the following problem is a relaxation of (2.3)

$$\min \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \left[\frac{a_{ij}}{\lambda_i}\right] x_j \leq \left[\frac{b_i}{\lambda_i}\right], \quad i = 1,\ldots,m \qquad (2.4)$$

$$x_j \text{ non-negative integer, } j = 1,\ldots,n,$$

where $[t]$ is the largest integer less than or equal to t. The constructive proof that (2.4) is a relaxation of (2.3) is exactly the same as the construction of a Gomory cut in the all integer method [4] and it will therefore be omitted.

It is also possible to restrict problem (2.3). Let

$$\langle t \rangle = \begin{cases} t & \text{if t is integer} \\ [t]+1 & \text{if t is not integer.} \end{cases}$$

Then the following problem is a restriction of (2.3)

$$\min \sum_{j=1}^{n} c_j x_j$$

$$\text{s.t.} \quad \sum_{j=1}^{n} \left\langle \frac{a_{ij}}{\lambda_i} \right\rangle x_j \leq \left\langle \frac{b_i}{\lambda_i} \right\rangle - 1 \qquad i = 1,\ldots,m \qquad (2.5)$$

$$x_j \text{ non-negative integer, } j = 1,\ldots,n.$$

If $\lambda_i | b_i$, then the right hand side of constraint i in (2.5) can be relaxed to $\langle b_i/\lambda_i \rangle = b_i/\lambda_i$.

LEMMA 1: Problem (2.5) is a restriction of problem (2.3).

PROOF: We want to show that x' feasible in (2.5) implies x' feasible in (2.3).

Consider a generic row i. Let

$$\left\langle \frac{a_{ij}}{\lambda_i} \right\rangle = \frac{a_{ij}}{\lambda_i} - f_{ij}, \qquad j = 1,\ldots,n,$$

and

$$\left\langle \frac{b_i}{\lambda_i} \right\rangle = \frac{b_i}{\lambda_i} - f_i$$

where $0 \leq f_{ij} < 1$, $j = 1,\ldots,n$, $0 \leq f_i < 1$.

Then

$$\sum_{j=1}^{n} \frac{a_{ij}}{\lambda_i} x'_j - \frac{b_i}{\lambda_i}$$

$$= \sum_{j=1}^{n} \left\langle \frac{a_{ij}}{\lambda_i} \right\rangle x'_j - \left\langle \frac{b_i}{\lambda_i} \right\rangle + f_i - \sum_{j=1}^{n} f_{ij} x'_j$$

$$< \sum_{j=1}^{n} \left\langle \frac{a_{ij}}{\lambda_i} \right\rangle x'_j - \left\langle \frac{b_i}{\lambda_i} \right\rangle + 1 \leq 0,$$

where the first (strict) inequality follows from the properties

$f_i < 1$ and $\sum_{j=1}^{n} f_{ij} x'_{ij} \geq 0$, for feasible x'_j and the second inequality

follows from the feasibility of x'_j. This completes the proof.

The construction of equivalent IP problems with coefficients smaller in magnitude is very difficult. In fact, an equivalent formulation of (2.3) with greatly reduced coefficients will generally not exist. Nevertheless,

sufficient conditions guaranteeing equivalence can be stated. These conditions will rarely be met but they may prove useful in selecting the divisors in the general case.

COROLLARY 1: The relaxed constraint $\sum_{j=1}^{n} \left[\frac{a_{ij}}{\lambda_i}\right] x_j \leq \left[\frac{b_i}{\lambda_i}\right]$ is an equivalent constraint if there exists a θ_i such that $\left\langle \frac{a_{ij}}{\theta_i} \right\rangle = \left[\frac{a_{ij}}{\lambda_i}\right]$, $j = 1,\ldots n$, and $\left\langle \frac{b_i}{\theta_i} \right\rangle - 1 = \left[\frac{b_i}{\lambda_i}\right]$.

Before we discuss the details of relaxation and restriction of IP problems, there are two points to be made. First, relaxation or restriction of a given IP problem may not only be desirable for reasons of efficiency, but it may also be necessary if group theoretic methods are to be applied. The size of the groups derived from (2.3) are equal to the magnitude of the determinants of the m x m non-singular matrices (bases) made up of the columns of (A,I), with particular interest focused on the determinant of the optimal LP basis. Computational experience has indicated that IP problems with basis determinants greater than 100,000 in magnitude are not very amenable to computation using group theoretic methods, and determinants less than 10,000 in magnitude are definitely preferred. Relaxation and restriction methods have been seen in practice to lower the magnitude of basis determinants.

Second, our current thinking is that relaxation is preferred over restriction because the transformation of an IP problem to a group optimization problem is also a relaxation. The composition of two relaxation is, of course, a relaxation. Thus, the lower bounds derived

from the optimal group problem solution for the relaxed problem is a lower bound on the original problem as well. Moreover, a feasible solution to the relaxed problem is optimal in the original problem if it is feasible. Thus, it is only necessary to test for feasibility in the original problem those solutions derived from group optimization problems that are feasible in the relaxed problem. The remainder of the paper will be concerned exclusively with relaxation procedures.

Finally, we remark that Kelley's dual cutting plane algorithm for the convex continuous programming problem is a relaxation method which converts the convex problem to a linear problem (see reference 8). Thus, it is clear that group theoretic methods and other IP relaxation methods can be combined with Kelley's algorithm to provide an algorithm for the convex IP problem. Kelley discusses the relationship of his algorithm to Gomory's cutting plane algorithm in reference 8, and we will onit further details here.

3. Relaxation of optimal bases

There is no single objective criterion for selecting the specific rational $\lambda_i \geq 1$, i=1,...,m, to use in constructing the relaxed IP problem (2.4). Consequently, the procedure for determining the λ_i must depend to a certain extent on heuristics derived from computational experience. These are discussed briefly in section 6 where computational experience is given. The task is made easier, however, by the fact that the λ_i can be chosen after problem (2.3) has been solved as an LP; that is, after an m x m optimal LP basis B made up of m linearly independent columns of (A,I) has been selected. For convenience, assume (A,I) = (N,B).

Let $B = (\gamma_{ij})$, let $D = |\det B|$, and let $B^{-1} = (\beta_{ij}/D)$; where γ_{ij} and β_{ij} are integers. Suppose a trial divisor $\lambda_r > 1$ is selected for row r. The rationale for selecting λ_r is that the relaxed matrix derived from B be non-singular. In so doing, the search for an optimal non-basic correction using group optimization problems (see [6]) is unchanged except for the coefficients in the canonical system over which the search is done.

Thus, we require that the matrix

$$R = \begin{pmatrix} \gamma_{11} & \cdots & \gamma_{1m} \\ \vdots & & \vdots \\ \gamma_{r-1,1} & \cdots & \gamma_{r-1,m} \\ \left[\frac{\gamma_{r,1}}{\lambda_r}\right] & \cdots & \left[\frac{\gamma_{r,m}}{\lambda_r}\right] \\ \gamma_{r+1,1} & \cdots & \gamma_{r+1,m} \\ \vdots & & \vdots \\ \gamma_{m,1} & \cdots & \gamma_{m,m} \end{pmatrix}$$

be non-singular. To see what condition ensures that $|\det R| \neq 0$, notice that RB^{-1} is an identity matrix except for row r which is of the form $(\xi_1,\ldots,\xi_m)$ where

$$\xi_i = \sum_{j=1}^{m} \left[\frac{\gamma_{r,j}}{\lambda_r}\right] \frac{\beta_{ji}}{D}, \qquad i=1,\ldots,m. \tag{3.2}$$

Clearly, since $\det RB^{-1} = \det R \det B^{-1} = \xi_r$, we have

$$|\det R| = D|\xi_r| = \left| \sum_{j=1}^{m} \left[\frac{\gamma_{rj}}{\lambda_r} \right] \beta_{jr} \right|, \tag{3.3}$$

and we require $\lambda_r \geq 1$ such that $\xi_r \neq 0$.

We note in passing that if the r^{th} slack variable is basic, then we must take $\lambda_r = 1$ because $\gamma_{rr} = \beta_{rr} = 1$, $\beta_{jr} = 0$, $j \neq r$. The inability to round a row for which the slack is basic in consistent with our goal of reducing the determinant. This is because a unimodular submatrix of B has no effect on the implied group structure of B.

Successive relaxations of B can be performed. Let $R_k = (\gamma_{ij}^k)$ be the m x m non-singular matrix obtained after k iterations of the relaxation procedure, and let $R_k^{-1} = \left(\frac{\beta_{ij}^k}{D_k} \right)$, where $D_k = |\det R_k|$. The new inverse R_{k+1}^{-1} can be computed from the inverse R_k^{-1} by the rule

$$R_{k+1}^{-1} = R_k^{-1} E_{k+1}$$

where E_{k+1} is an elementary transformation matrix. In particular, if rounding row r by λ_r is the $(k+1)^{st}$ relaxation, then E_{k+1} is an m x m identity matrix except for row r which is

$$(-\xi_1^k/\xi_r^k, \ldots, -\xi_{r-1}^k/\xi_r^k,\ D_k/\xi_r^k,\ -\xi_{r+1}^k/\xi_r^k, \ldots,\ -\xi_m/\xi_r)$$

where

$$\xi_i^k = \sum_{j=1}^{m} \left[\frac{\gamma_{rj}^k}{\lambda_r} \right] \frac{\beta_{ji}^k}{D_k} \quad .$$

Upon termination of the basis relaxation procedure, it is necessary to relax the matrix N of non-basic activities, and also relax the right hand side b. The divisors are, of course, the same as those used to relax the basis, and the relaxations should be done in the same order.[1]

We now turn to a study of the determinant of the relaxed basis R as a function of the divisor λ_r. Define the determinant function

$$\delta(\lambda) = \sum_{j=1}^{m} \left[\frac{\gamma_j}{\lambda}\right] \frac{\beta_j}{D} \qquad (3.4)$$

which is the expression (3.2) for i=r, where the index r on the coefficients γ_{rj}, β_{jr} and λ_r have been suppressed. Clearly, $\delta(\lambda)$ will not in general be a monotonic function of λ, and the determinant of the relaxed basis may vary erratically as λ is increased. Although it is possible to give some results on the behavior of $\delta(\lambda)$, we omit such an analysis because it is straightforward.

We remark tangentially that the choice of a particular optimal LP basis when there is large scale LP degeneracy should not always be left to the simplex method. Computational experience on a limited number of test problems has suggested the following procedure. Let the simplex method find an optimal solution and some set of optimal LP reduced costs. If the optimal basis determinant is too large, take the optimal basic activities which are at a positive level and pivot them into an identity matrix. The result is a basis with many basic slacks and this basic solution is the same as the previously obtained optimal basic solution. The shadow prices relative to this basis may not be dual feasible, but the previously obtained optimal LP reduced costs are non-negative and represent a consistent set of cost coefficients to use in the subsequent analysis.

---

[1] Slack variable activities are not relaxed. This follows from our discussion in section 2.

We cite the following example. A scheduling problem with 259 rows was found to have an optimal basis determinant at the end of the simplex method of approximately 103,000. There was, however, large scale degeneracy, and the procedure outlined above led to an optimal basis corresponding to the same solution with determinant equal to 8.

4. Relaxation and Generalized Lagrange Multipliers

In this section, we give a reformulation of the IP problem (2.3). The reformulation includes an identification of the inaccuracies due to rounding the optimal LP basis. As a result, it is possible to use generalized Lagrange multipliers to "price out" the inaccuracies. These multipliers can be found by the generalized LP procedure of Brooks and Geoffrion [2]. The discussion of the properties and algorithmic uses of the multipliers is brief in order not to duplicate the discussion in reference 10 about similar problems and multipliers.

Recall that the coefficient matrix (A,I) of (2.3) is reordered and partitioned as (N,B) where B is an optimal LP basis to (2.3). Let (v,y) by the corresponding partition of the variables into non-basis and basic variables, respectively. Let the m x m non-singular matrix R be a relaxation of B, and let Q be the corresponding relaxation of N. Let q_j, j=1,...,n, be the columns of Q. Finally, let d be the corresponding relaxation of b, and let g_R be the canonical homomorphism which maps the group M(I) onto the factor group $\mathcal{G} = M(I)/M(R)$.

The following problem can easily be shown to be equivalent to (2.3)

$$\min \sum_{j=1}^{n} \bar{c}_j v_j \tag{4.1a}$$

$$\text{s.t.} \quad \bar{N}v + Iy = \bar{b} \tag{4.1b}$$

$$\bar{Q}v + Iy = \bar{d} \tag{4.1c}$$

$$\sum_{j=1}^{n} \rho_j v_j = \rho_0 \tag{4.1d}$$

$$v_j \text{ non-negative integer}, \quad j=1,\ldots,n \tag{4.1e}$$

$$y \geq 0 \tag{4.1f}$$

where $\bar{c}_j = c_j - c_B B^{-1} a_j \geq 0$, $\bar{N} = B^{-1}N$, $\bar{b} = B^{-1}b$, $\bar{Q} = R^{-1}Q$, $\bar{d} = R^{-1}d$, $\rho_j = g_R(q_j)$, $j=1,\ldots,n$, $\rho_0 = g_R(d)$. To see that problem (4.1) is equivalent to (2.3), note that conditions (4.1d) and (4.1c) ensure that $y = \bar{d} - \bar{Q}v$ is integer. Condition (4.1f) ensures that $y \geq 0$. Condition (4.1b) ensures that the non-negative integer solution (v,y) is feasible in the original problem. On the other hand, any solution that is feasible in (2.3) is feasible in (4.1).

It is convenient to eliminate explicit mention of y in (4.1) and hence we rewrite it as

$$\min \sum_{j=1}^{n} \bar{c}_j v_j \tag{4.2a}$$

$$(\bar{N} - \bar{Q})v = \bar{b} - \bar{d} \tag{4.2b}$$

$$\bar{Q}v \leq \bar{d} \tag{4.2c}$$

$$\sum_{j=1}^{n} \rho_j v_j = \rho_0 \tag{4.2d}$$

$$v_j \text{ non-negative integer}, \quad j=1,\ldots,n \tag{4.2e}$$

Let $V = \{v \mid v \text{ satisfies (4.2d), (4.2e)}\}$, and let $[V]$ denote the convex hull of V; $[V]$ is an unbounded convex polyhedron with integer extreme points and its faces constitute the strong cuts for the rounded problem (see reference 5). Thus we consider the LP problem

$$\min \bar{c}v \tag{4.3a}$$

$$\text{s.t.} \quad (\bar{N} - \bar{Q})v = \bar{b} - \bar{d} \tag{4.3b}$$

$$\bar{Q}v \leq \bar{d} \tag{4.3c}$$

$$v \in [V] \tag{4.3d}$$

We solve problem (4.3) by generalized LP (this is the suggestion of Brooks and Geoffrion [2]) rather than generate all the faces of $[V]$. The details for doing this will be omitted because they are given for a very similar problem in reference 10. We simply mention that a group optimization problem defined over $\mathcal{R}$ is used to generate solutions for a master LP problem involving the constraints (4.3b) and (4.3c). This problem is of the form

$$\min (\bar{c} + \theta^0(\bar{N} - \bar{Q}) + \pi^0\bar{Q})v \tag{4.4}$$

$$\text{s.t.} \quad v \in V$$

where θ^0 and π^0 are the (generalized) LP multipliers from the master. The vector θ^0 is unconstrained in sign, and the vector $\pi^0 \geq 0$. Clearly, θ^0 "prices out" the inaccuracies due to rounding, while π^0 "prices out" the resources $\bar{d}$ available for finding an optimal correction v in the rounded problem.

Let v^0 denote an optimal solution in (4.4). The cut

$$(\bar{c} + \theta^0(\bar{N} - \bar{Q}) + \pi^0\bar{Q})v \geq (\bar{c} + \theta^0(\bar{N} - \bar{Q}) + \pi^0\bar{Q})v^0 \qquad (4.5)$$

is a valid Gomory cut and can be written with respect to any multipliers θ^0 and $\pi^0 \geq 0$. However, a particularly strong cut is obtained if θ^0 and π^0 are the optimal shadow prices in the LP problem (4.3).

5. Relaxation and Bender's Method for Mixed Integer Programming

In this section, we exploit the relaxation theory developed above to combine group theoretic methods for pure IP with Bender's mixed IP algorithm [1]. We assume that the reader is familiar with Bender's algorithm, and consequently, our development will concentrate on the features of the algorithm proposed here that are new.

The mixed IP problem is written as

$$\begin{aligned} &\min\ cv + qw \\ \text{s.t.}\quad &Av + Pw \geq b \qquad (5.1) \\ &v \geq 0,\ w \in W \end{aligned}$$

where

$$W = \{w \mid Rw = r, \ w \text{ non-negative integer}\}.$$

The vector c is 1 x n_1, the vector q is 1 x n_2, the vector b is m_1 x 1, and the vector r is m_2 x 1. All coefficients are integers.

Define the convex polyhedron

$$U = \{u \mid uP \leq q, \ u \geq 0\} \tag{5.2}$$

and let u^t, $t \in T$, be the extreme points of U, and let u^s, $s \in S$, be a complete set of extreme rays of U satisfying $u^sP \leq 0$, $u^s \geq 0$. Bender's algorithm assumes that u^t, $t \in T'$, u^s, $s \in S'$ are given where $T' \subseteq T$, $S' \subseteq S$. Bender's algorithm begins with the problem

$$\min w_0 \tag{5.3a}$$

$$\text{s.t.} \quad w_0 - (q - u^tP)w \geq u^tb, \ t \in T' \tag{5.3b}$$

$$-u^sPw \geq -u^sb, \ s \in S' \tag{5.3c}$$

$$w \in W \tag{5.3d}$$

Problem (5.3) is a mixed IP problem where the continuous variables are w_0 and the slacks on the constraints (5.3b) and (5.3c). Our version of the algorithm begins instead with the following relaxation of (5.3).

$$\min w_0$$

$$\text{s.t.} \quad w_0 - \lceil (q - u^tP) \rceil w \geq \lceil u^tb \rceil + 1, \ t \in T'$$

$$- \lceil u^sP \rceil w \geq -\lceil u^sb \rceil, \ s \in S' \tag{5.4}$$

$$w \in W$$

$$w_0 \text{ integer}$$

Notice that (5.4) is now a pure IP problem because w_0 and the slacks will naturally take on integer values.

Suppose (5.4) is solved using the IP algorithm of [6] and let w_0^*, w^* be the optimal solution. The next step is to solve the LP problem

$$\begin{aligned} & \min cv \\ \text{s.t.} \quad & Av \leq b - Pw^* \\ & v \geq 0 \end{aligned} \tag{5.5}$$

If (5.5) has an unbounded solution, then (5.1) has an unbounded solution.

If (5.5) has a (bounded) optimal solution, say v^*, with optimal dual variables u^*, then there are three possibilities. First, if $w_0^* \geq qw^* + cv^* = (q - u^*P)w^* + u^*b$, then (w^*, v^*) is an optimal solution to (3.1). On the other hand, if $w_0^* < (q - u^*P)w^* + u^*b$, then a constraint is added to (5.4) and (5.4) is reoptimized. There are two cases to consider when a constraint is added. The ordinary case is the addition of the constraint

$$w_0 \geq \lceil (q - u^*P) \rceil w + \lceil u^*b \rceil + 1$$

to (5.4). The exceptional case arises when $w_0^* < (q - u^*P)w^* + u^*b$, but $w_0^* \geq \lceil (q - u^*P) \rceil w^* + \lceil u^*b \rceil + 1$ implying that a more restrictive relaxation of the desired constraint is required. We will discuss below the steps to be taken in the latter case.

Finally, it may be that (5.5) is infeasible, a condition that implies the existence of an extreme ray u^s of U with the property that $u^s(b - Pw^*) > 0$. In this case, add the constraint

$$-\lfloor u^s P \rfloor w \geq -\lfloor u^s b \rfloor$$

to (5.1) if this constraint is binding. If this constraint is not binding, i.e., $-\lfloor u^s P \rfloor w^* \geq -\lfloor u^s b \rfloor$, then once again a more restrictive relaxation of $-u^s Pw \geq u^s b$ is required.

Suppose now that the relaxed problem (5.4) has yielded an optimal solution w_0^*, w^* such that $w_0^* < (q - u^*P)w^* + u^*b$, but $w_0^* \geq \lfloor (q - u^*P) \rfloor w^* + \lfloor u^*b \rfloor + 1$. In this case, a variety of options are open to us, and it is difficult to anticipate which ones will be preferred for computation. Several procedures are suggested in [12]. The approach we take here is to give one simple extension, and await further computational experience. First, we drop all the constraints (5.3b) in favor of the single constraint $w_0 \geq (q - u^*P)w + u^*b$. Since the data must be given in integers, we rewrite this as $D^*w_0 \geq D^*(q - u^*P)w + D^*u^*b$, where D^* is the absolute value of the determinant of the optimal LP basis in (5.5) which yielded u^*. Fina we substitute the variable z_0 for D^*w_0; z_0 will be an integer variable.

The second change is to replace the constraints (5.3c) by $-D_s u^s P w \geq -D_s u^s b$, where D_s is once again the corresponding magnitude of the determinant. (This change is, of course, required when $\lfloor u^s P \rfloor w^* < \lfloor u^s b \rfloor$, but the LP problem (5.5) is infeasible and u^s is an extreme ray of U with the property $u^s(b - Pw^*) > 0$).

Thus, problem (5.3) becomes

$$\begin{aligned} & \min\ z_0 \\ \text{s.t.}\quad & z_0 - D^*(q - u^*P)w \geq D^*u^*b \\ & -D_s u^s P w \geq -D_s u^s b,\ s \in S' \\ & z_0 \text{ integer},\ w \in W \end{aligned} \tag{5.6}$$

Problem (5.6) is a pure integer problem. It is solved first as an LP problem, and the optimal LP basis is examined. If the determinant is too large in magnitude, then the relaxation procedures of section 3 are appropriate. Finally, the group theoretic methods are applied and an optimal z_0^*, w^* obtained. This solution is tested as before for optimality in LP (5.5). If a new constraint on z_0 is indicated, then the new constraint should replace the old one in (5.6). A test should be made to see that no extreme point of the polyhedron U given by (5.2) is repeated.

There are many algorithmic details to be made precise when the method outlined above is encoded. As we have said, some of these will depend upon computational experience not yet realized. There are, however, two ideas worth discussing. First, it is important that the first vector u of shadow prices in (5.4) be meaningful, and it may be worthwhile investing a relatively large amount of computational effort to obtain them. An obvious possibility is to solve (5.1) as an LP and use the shadow prices on the constraints $Av + Pw \geq b$ as the first set in (5.4). For some problems, the LP shadow prices may be somewhat inaccurate in their reflection of system economies and diseconomies.

An alternative is to use Gomory's cutting plane method for mixed integer programming until it begins to slow down.

A second computational issue in using Bender's method is how to efficiently update the solution of problem (5.4) or problem (5.6) when a new constraint is added. A number of good suggestions are made in the article by Lemke and Spilberg [9]. We simply mention that it is easy to add side constraints to an IP problem previously solved by group theoretic methods. In particular, the previously optimal LP basis for (5.4) or (5.6) can be used as the dual feasible basis from which the group problems are induced.

6. Computational Experience

Our computational experience thus far with relaxation methods had been with ad hoc procedures designed for specific applications. As such, the experience is difficult to organize, but it has been instructive and we feel it is worthwhile discussing it here. The indication in all cases is that relaxation after obtaining an optimal LP basis will produce much better procedures.

One of the first experiments with relaxation was on a minimal cost scheduling problem which consisted of 66 rows and more than 400 activities (columns). Due to core storage limitations at the time the problem was run, it was not possible to consider all of the activities, and a subset of 176 plus a full set of slacks was selected. Included were all the columns of an optimal LP basis which had a determinant of 4860. This determinant was not too large for computation with the unconstrained group algorithm (see reference 6), but it was decided that some relaxation was preferred. In particular, there was a less than or equal to inequality with many non-zero (integer) coefficients between 13 and 35, and a right hand side of 308.

Relaxation of the original data (i.e., before running the simplex algorithm) of this single constraint using the divisor $\lambda = 2$ yielded an optimal LP basis with a determinant of 408, and relaxation with $\lambda = 4$ yielded a determinant of 66. The latter relaxed IP problem was solved in approximately 90 seconds on a UNIVAC 1108 computer. One feasible and nearly optimal solution to the relaxed problem used up 330 units of the relaxed constraint's resource, and the optimal solution used up 341 units. These solutions were not entirely satisfactory, but further useful computation was not done.

A heuristic using generalized Lagrange multipliers was suggested for this problem. The optimal LP shadow price on the constraint that was relaxed was \$3.15, and the total cost equal to \$3.15 x 308 due to that row was approximately 70% of the minimal LP cost for the problem. The experience discussed above indicated that this single constraint was making the IP problem an order of magnitude more difficult. Thus, the heuristic of adding the constraint to the objective function with a "cost" of \$3.50, say, thereby eliminating it from the problem would have been a good experiement to try.

Another problem for which relaxation was required was a brand allocation problem which arises in the scheduling of spot commercials on TV. A large company buys many spots as a package, and then must decide how to equitably and profitably allocate the spots to its different brands, and it must do this subject to a variety of logical constraints. A test problem consisted of 86 constraints and 148 zero-one variables. Additional upper bound constraints on the variables were handled implicitly. Most of the constraints were logical constraints with non-zero coefficients equal to 1. There were, however, four constraints with large coefficients of three and four place integers, each required to be satisfied to within 5% of the stated right hand side. Each of these constraints was written as a less than or equal to inequality with an upper bound of 10% on the slack variable.

No relaxation of these four constraints was done for the initial run, and the optimal LP basis had a determinant on the order of 10^{39}. The constraints were then rounded until all of the non-zero coefficients were one digit. The determinant of the optimal basis of the relaxed problem was then 380. However, the inaccuracies due to relaxation allowed one constraint to be violated by more than 10% above the target value, and another was below the target value by more than 5%. We then employed heuristics to get the solution within plus or minus 5% on all four constraints. The heuristics consisted of small adjustments in the right hand sides and reductions in the upper bounds on the slacks. The results are summarized in Table I below. The solution for run 5 satisfied the four relaxed constraints, but an earlier solution which slightly violated these constraints was preferred.

Finally, we discuss a class of problems for which relaxation of the original data was particularly difficult. This is the media selection problem in which various TV and radio programs, and magazines are selected for the advertising campaign of a single product. The objective function consists of effectiveness coefficients for each media decision variable. Included among the constraints are yearly and monthly budget constraints whose coefficients were given to four places. This problem resembles some of the capital budgeting problems discussed by Weingartner in reference 11.

There were two difficulties caused by relaxation of the original data. The first was the disparity between the cost per unit of the various media. A 30 second radio spot may cost $500, while a 60 second nighttime network TV spot may cost $50,000. Sufficient relaxation of a budget constraint with coefficients differing in magnitude by 100 is impossible without eliminating the smaller coefficients. A solution to this difficulty was to aggregate the low cost per unit media in order to make them more comparable on a cost per unit basis.

A second difficulty with relaxation of the original data was the inaccuracy produced in the benefits per unit of closely competing media. Relaxation after solving the LP is definitely preferred for these problems. Inaccuracies in non-basic variables is a lower order effect because non-basic variables are used more sparingly. One heuristic we used was to solve the original problem as an LP and thereby identify the optimal basic activities. Accuracy in describing the activities of these variables was favored in subsequent runs on relaxed data. Possible deleterious effects of inaccuracies among the non-basics was reduced by placing tight upper bounds on the non-basics.

Table I

	Run 1	Run 2	Run 3	Run 4	Run 5
D	380	168	380	200	280
Optimal LP value	220983	224260	227160	224985	220258
Optimal IP value	219600	223380	226280	224349	217890
LP time	18.37	27.08	18.38	12.82	18.38
Group representation time	1.58	1.30	1.59	1.52	1.46
Unconstrained group optimization time	1.40	0.89	1.39	0.95	1.93
Search time	23.95	3.17	2.09	2.59	43.08
Total run time	59.0	44.0	35.0	29.0	78.0

Times are in seconds on a UNIVAC 1108 computer.

References

1. Benders, J.F., "Partitioning Procedures for Solving Mixed Variables Programming Problems," Numerische Mathematik 4, 238-252 (1962).

2. Brooks, R. and A. Geoffrion, "Finding Everett's Lagrange Multipliers by Linear Programming," Operations Research 14, 1149-1153 (1966).

3. Geoffrion, A.M., "Relaxation and the Dual Method in Mathematical Programming," Western Management Science Institute Report No. 135, UCLA, (March 1968).

4. Gomory, R.E., "All Integer Programming Algorithm," pp. 193-206 in J.F. Muth & G.L. Thompson (eds.), Industrial Scheduling, Prentice-Hall (1963).

5. Gomory, R.E., "Some Polyhedra Realted to Combinatorial Problems," IBM Research Report RC2145, (July 1968).

6. Gorry, G.A. and Shapiro, J.F., "An Adaptive Group Theoretic Algorithm for Integer Programming Problems," Technical Report No. 39, Operations Research Center, M.I.T., May 31, 1968. (Revised May 1969). To appear in Management Science.

7. Hu, T.C., "Integer Programming and Network Flows," Addison-Wesley (1969).

8. Kelley, J.E., Jr., "The Cutting Plane Method for Solving Convex Programs", SIAM J. on Appl. Math., 8, 703-712 (1960).

9. Lemke, C.E. and Spielberg, K., "Direct Search Algorithms for Zero-One and Mixed-Integer Programming," Operations Research, 15, 892-914 (1967).

10. Shapiro, J.F., "Generalized Lagrange Multipliers in Integer Programming," Sloan School of Management, M.I.T., Working Paper 350-68, March 1969. (Revised January 1970).

11. Weingartner, H.M., "Capital Budgeting of Interrelated Projects: Survey and Synthesis," Management Science, 12, 485-516 (1966).

12. Wolsey, L.A., "Mixed Integer Programming: Discretization and the Group Theoretic Approach," Ph.D. Dissertation, M.I.T., (1969).

Date Due

NOV 3 '76
AUG 29 '83
FEB 14 '76
DEC 07 '78
MAR 3 '77
MAR 28 '77
APR 22 '77
NOV 13 '79
APR 3 '78

Lib-26-67

MIT LIBRARIES
3 9080 003 875 660
454-70
MIT LIBRARIES
3 9080 003 906 663
456-70
MIT LIBRARIES
3 9080 003 906 630
HD2
.M41
Nos.1
Nos.46
MIT LIBRARIES
457-70
3 9080 003 906 606
MIT LIBRARIES
458-70
3 9080 003 906 671
MIT LIBRARIES
459-70
3 9080 003 875 652
MIT LIBRARIES
460-70
3 9080 003 906 622
MIT LIBRARIES
461-70
3 9080 003 875 611
MIT LIBRARIES
462-70
3 9080 003 875 637
MIT LIBRARIES
463-70
3 9080 003 875 678
MIT LIBRARIES
464-70
3 9080 003 875 645
MIT LIBRARIES
465-70
3 9080 003 906 580